Reliability and Responsiveness of Disablement Measures Following Acute Ankle Sprains Among Athletes[1]

By: Rick W. Wilson, PhD, PT[2], Joe H. Gieck, EdD, PT, ATC[3], Bruce M. Gansneder, PhD[4], David H. Perrin, PhD, ATC[5], Ethan N. Saliba, PhD, PT, ATC, SCS[6], and Frank C. McCue, Ill, MD[7]

Wilson, R.W., Gieck, J.H., Gansneder, B.M., Perrin, D.H., Saliba, E.N., & McCue, F.C. (1998). Reliability and responsiveness of disablement measures following acute ankle sprains among athletes. Journal of Orthopedic and Sports Physical Therapy, 27:348-355.

*****Note: Figures may be missing from this format of the document**

Abstract:

The minimum amount of change which represents clinical improvement following ankle sprains is unknown. This study considers the usefulness of physiological and behavioral measures commonly employed for this purpose in sports rehabilitation settings. Thirteen collegiate athletes of both genders were measured at approximately 3 and 10 days post-grade I or II ankle sprain. Volumetric displacement and sagittal plane ankle range of motion measures were used as impairment indicators. Motor ability scores (activity count) and a perceived athletic ability measure (visual analog scale) were used to indicate functional limitations. Volumetric displacement and both functional limitation measures demonstrated responsiveness to change between two occasions of measurement separated by 1 week. Observed changes in range of motion deficits could not be distinguished from measurement error. The results of this study suggest that behavioral measures of motor activity and perceived athletic ability may be at least as useful as physical measures of organic dysfunction for assessing clinical improvement following acute ankle sprains among athletes.

Article:

Rehabilitation professionals use a variety of assessment tools to arrive at clinical decisions following ankle sprains among athletes. Investigators have used volumetry (1,5,6,10,17,19,21,27), goniometry (1,5,6,17,19,21), observed locomotor ability

---

[1] 1996 Breg Excellence in Research Award paper. Presented at the 17th Annual Physician-Therapist Team Concept Conference, Oak Brook, IL, October 11-13, 1996.

[2] Assistant Professor, Department of Physical Therapy, Moore Hall, Room 308, Western Carolina University, Cullowhee, NC 28723

[3] Professor, Department of Human Services, Curry School of Education, University of Virginia, Charlottesville, VA; Associate Professor, Department of Orthopedics, University of Virginia Medical School, University of Virginia, Charlottesville, VA

[4] Professor, Department of Educational Studies, Curry School of Education, University of Virginia, Charlottesville, VA

[5] Professor and Director, Program in Health and Physical Education, Curry School of Education, University of Virginia, Charlottesville, VA

[6] Assistant Professor, Department of Orthopedics, University of Virginia Medical School, University of Virginia, Charlottesville, VA

[7] Alfred R. Shands Professor of Orthopedics and Plastic Surgery of the Hand and Director, Division of Sports Medicine and Hand Surgery, Department of Orthopedics, University of Virginia Medical School, University of Virginia, Charlottesville, VA

(5,6,10,19,21,27), and self-reported functional ability (1,6,10,17) as dependent measures in ankle sprain studies designed to demonstrate the efficacy of various treatments. However, the usefulness of these measures for the purpose of monitoring treatment response has not been established.

One reason for this situation may be the preponderance of between- group research designs in the rehabilitation literature. All of the ankle sprain studies cited above were designed to detect differences between various treatment and control groups. However, researchers who failed to detect treatment effects between groups were unable to determine whether treatment was truly ineffective or the dependent measure was simply unresponsive to the effects of the treatment.

The property of responsiveness refers to a measure's ability to detect changes from one occasion to an-

Sport	Males	Females
Baseball	1	0
Basketball	2	0
Field hockey	0	1
Football	1	0
Gymnastics	0	1
Lacrosse	4	0
Soccer	1	2
Softball	0	1
Track	3	2
Volleyball	0	1
Wrestling	1	0
Total	13	8

TABLE 1. *Subject characteristics: Sport by gender data (N = 21).*

other which exceed those expected by chance (9,16). Although several ankle sprain studies have included repeated measures of the dependent variables within groups of subjects (1,5,6,14,17,19,20,27), none have compared the magnitude of changes occurring within subjects due to treatment to variability observed during a stable baseline period.

A fascination with the statistical significance of hypothesis tests may also have contributed to the dearth of responsiveness data in the rehabilitation literature. Statistical significance is influenced by sample size, and minor changes in average values may be found to be statistically significant if sufficiently large samples are used. Conversely, statistically significant research findings may appear trivial and irrelevant to clinicians who are interested in achieving meaningful clinical responses to treatment in individual patients (15). A more useful approach to statistical analysis of research results considers the size of the differences occurring in response to treatment or the passage of time. However, results based solely on average differences between or within groups provide no information about individual subjects' responses to treatment, and minimum meaningful change values for individuals have not been determined for even the most commonly used clinical measures (9). Responsiveness data can be used to estimate the smallest amount of change in a measure which may be considered to represent a true improvement in individual subjects (2,9,11).

Lacking proven measures for monitoring treatment response, many researchers and clinicians exhibit a selection bias when choosing clinical assessment tools. Physical measures of impairment are generally regarded as trustworthy, while behavioral measures are rarely quantified or given adequate consideration in either clinical decision-making or research (3). This reluctance to trust and use behavioral measures may be an expression of the belief that observational and self-reported information is somehow unscientific (22). However, whether an instrument is a behavioral observation or self-report rather than a physical measure is not necessarily a relevant scientific distinction (13,22,23). A more rational approach to instrument selection is based on an assessment measure's demonstrated usefulness for a particular purpose.

Our primary objective in this study was to evaluate and compare the usefulness of range of motion deficits (ROMLOSS), volumetric measurements of swelling (SWELLING), athletes' self-reports of perceived athletic ability (ATHABILITY), and observed motor activity scores (ACTIVITY) as indicators of clinical progress following ankle sprain. To achieve this goal, we examined the reliability and responsiveness of these four measures. Whether behavioral measures of functional limitation (ACTIVITY and ATHABILITY) were more or less responsive than measures of physical impairment (SWELLING and ROMLOSS) was a theoretical issue which became a secondary research question within the context of this study.

METHODS

Study Design

A prospective, multivariate, within-subjects design was employed to observe the stability and responsiveness of clinical measures during the early clinical rehabilitation period following acute ankle sprain. The clinical rehabilitation period was defined as the interval between the date of injury and return to athletic participation without limitations. Initial measurements were taken approximately 3 days (67.8 ± 15.2 hours) postinjury. Multiple measurements of each variable were taken by the same investigator during this baseline session to determine intraoccasion test-retest reliability and standard errors of measurement. Data from a second measurement session held approximately 1 week later (6.42 ± 1.25 days) were also recorded by the same investigator and analyzed for evidence of each measure's responsiveness to change.

Subjects

Twenty-four consecutive athletes with grade I or II ankle sprains volunteered for this study. Informed consent to participate was obtained according to guidelines issued by the Committee for the Protection of I Inman Subjects, University of Virginia, Charlottesville, VA. Three volunteers failed to appear for the initial measurement session. Data for these three subjects were deleted from the statistical analysis in a list-wise fashion. Sample characteristics for the remaining 21 athletes are displayed in Table 1.

Of the 21 subjects participating in the baseline (test-retest reliability) measurement session, five subsequently returned to full participation within a week and did not participate in further treatment or reevaluation sessions. Three additional subjects were lost to follow-up during breaks in the academic calendar or team travel periods. Data from these eight subjects were deleted from the responsiveness calculations in a list-wise manner.

Instruments and Measurements

All measurements were taken using the following procedures in the same order in which they appear.

Joint Swelling

Swelling of the affected ankle and foot was evaluated by the water displacement method. After removing socks and shoes, subjects were measured while sitting with the knee and hip flexed to approximately 90°. A thin coat of water was applied to the limb prior to immersion to minimize the amount of air trapped around leg hair. Volumetric displacement of both ankles was measured using a commercially available foot volumeter (Smith & Nephew-Rolyan, Menomonee Falls, WI) using water at a temperature of 80-85°C. Subjects were instructed to gently lower the limb into the water until the foot rested comfortably on the bottom of the volumeter with the heel and calf positioned to touch its rear wall. In this position, the lower extremity was submerged to a depth of approximately 25 cm. The displaced volume of water was discharged through an overflow spout, captured in a basin, and transferred to a 1,000-m1 graduated cylinder for measurement. The displaced volume of water was recorded to the nearest 5 ml for each of two consecutive trials.

Range of Motion

Subjects were placed in the prone position with the knee flexed to approximately 90°. Passive dorsiflexion and plantar flexion were recorded bilaterally. Specific anatomical landmarks employed were the lateral malleolus (axis of motion), lateral aspect of the fifth metatarsal head (distal arm), and an imaginary line between the lateral malleolus and the fibular head (proximal arm). Subjects were instructed to inform the investigator of any pain experienced during the measurement. If pain was experienced during dorsiflexion or plantar flexion, the indicated point of pain onset was used as the position of end range measurement. If no pain was encountered, passive range of motion end points were established using light manual overpressure at the fourth and fifth metatarsal heads. Total range of sagittal plane motion of the ankle and foot (maximum dorsiflexion to maximum plantar flexion) was recorded to the nearest degree during each of three consecutive measurements.

Motor Activity Score

At the beginning of the session, subjects were presented with the following list of closed chain tasks and received a demonstration of each task from the investigator. Following the demonstration, subjects were asked whether they could perform that activity comfortably. If the subjects answered yes, they were invited to perform the activity. All subjects were instructed to discontinue a task immediately if they experienced symptoms of ankle pain or instability. A dichotomous scoring system was employed for each task, with one point awarded for successful completion of a task and no points awarded if a task was not attempted or aborted. Subjects wore their usual athletic shoes during all activities, which were performed on a smooth concrete surface.

40-meter ambulation Subjects walked full weight bearing without crutches through two trials over a straight 40-m course.

40-meter run Subjects ran a straight 40-m course during each of two trials.

Figure-8 run Cones at each end of a 6-m course were designated as circling points. Subjects ran a figure-8 path around the cones during each of two trials. Each trial consisted of two laps of the course, approximately 24 m.

Single hop Subjects stood on the injured limb, then hopped as far as possible on two successive trials, landing on the same limb. A hop was considered completed if a subject stood on the affected limb, cleared the ground, and returned to the original unilateral full weight-bearing position without contacting the ground with the uninvolved extremity.

Cross-over hop This activity was performed over a 6-m course with a 15-cm wide marker on the ground along its length. Subjects hopped three consecutive times on the injured limb, crossing over the center strip marker with each hop.

Stairs hop Subjects hopped up and down a flight of 14 steps on the injured extremity. Each step was approximately 20 cm high.

Self-Reported Athletic Ability

A visual analog scale was used to assess the subjects' response to the question "Compared to normal, how would you rate your athletic ability today?" by making a single vertical mark on a 100-mm line situated between two polar descriptors ("Normal, no limitation" and "Severely limited"). Responses were converted to numerical scores by measuring the distance from the zero point ("Normal, no limitation") to the vertical mark. This distance was recorded to the nearest millimeter.

Data Analysis

Prior to analysis, the distributions of scores for all four measures on both occasions of measurement were examined for skewness and kurtosis using a modified Kolmogorov-Smirnov Z (Lilliefors) test for normality (SPSS for Windows, version 6.1., SPSS Inc., Chicago, IL). The distributions of scores observed on the second occasion of measurement (approximately day 10 postinjury), ROMLOSS, and ACTIVITY demonstrated skewness, kurtosis, or both.

Consequently, distribution-free statistics were used to analyze responsiveness to change in all four measures. Missing data were deleted from this analysis in a pair-wise manner.

Reliability

Intraoccasion test-retest reliability coefficients and standard errors of measurement were calculated from univariate repeated-measures analysis of variance values using Shrout and Fleiss' formula 2,1 (24). The test-retest reliability of the two derived measures, ROMLOSS and SWELLING, was also examined. ROMLOSS was computed by subtracting range of motion measurements taken on the injured ankle from measurements taken on the uninvolved ankle. SWELLING was computed by subtracting measures of volume displacement taken from the uninvolved ankle from similar measures taken from the injured one. Internal consistency (Cronbach's alpha) of the additive motor ability score was assessed using SPSS' Reliability (SPSS for Windows, version 6.1., SPSS Inc., Chicago, IL).

Responsiveness

Responsiveness of a measure depends on its ability to detect changes in scores which exceed those expected by chance. Reliable change coefficients were calculated for each measure using the method described by Christensen and Mendoza (2).

$$\text{Reliable change} = \frac{\Delta}{\sqrt{2 \times \text{SEM}^2}}$$

In this study, the numerator represents time-dependent change, the difference between scores obtained on two occasions of measurement separated by 1 week. Scores obtained from the initial trial on the first occasion of measurement and the second occasion of measurement were used to compute this difference for each of the four measures. The denominator reflects measurement error, the

	Item-Total Correlation	Alpha If Item Deleted
40-m walk	.629	.873
40-m run	.792	.855
Figure-8 run	.835	.850
Single hop	.815	.852
Cross-over hop	.815	.852
Stair hop	.689	.868
		Alpha = 0.90

TABLE 2. *Internal consistency (Cronbach's alpha) of the motor activity scale (N = 21).*

within-subjects' variability of scores obtained during the first occasion of measurement. We defined a minimum reliable change to be one which was at least 1.96 reliable change removed from baseline values. This method produces critical reliable change values for detecting changes in measures which are unlikely ($p < 0.05$) to occur due to chance (11). We chose to use ICC values from the preceding reliability study in place of the Pearson correlation coefficients (r) customarily used to calculate the standard error of measurement (SEM).

$$SEM = \sigma\sqrt{1 - r}$$

This method provides standard error of measurement estimates which are slightly more conservative than those derived using Pearson r values, since ICCs are sensitive to both random and systematic measurement errors.

Bonferonni corrections to four preplanned comparisons (Wilcoxon signed-rank tests for paired samples) were employed to test the hypotheses that ROMLOSS, SWELLING, ACTIVITY, and ATHABILITY were responsive to changes occurring between two measurement occasions separated by approximately 1 week. A corrected alpha of 0.0125 (two-tailed) was established for each test to maintain a study-wide type I error rate of 0.05.

RESULTS

Reliability/Internal Consistency

Internal consistency of the motor activity score was 0.90, indicating an acceptable degree of additivity among the dichotomously scored locomotor tasks (Table 2). Therefore, a motor activity score (ACTIVITY) was computed by summing the individual task scores for each subject. Intraoccasion test-retest reliability coefficients (ICC21) for the four predictor variables ranged from 0.86 to 1.0 (Table 3).

Responsiveness

Within-group changes occurring between postinjury clays 3 and 10 were detected with three of the indicator variables: SWELLING ($p = 0.0015$) (Figure 1), ACTIVITY ($p = 0.0015$) (Figure 2), and ATHABILITY ($p = 0.0015$) (Figure 3). However, decreasing range of motion deficits occurring over 1 week could not be distinguished from measurement error ($p = 0.165$) (Figure 4). Individual subjects' change scores were then examined for responsiveness using

	$ICC_{2,1}$	SEM	Minimum Reliable Change	Mean Observed Change
ROMLOSS	0.88	4.1°	11.3°	4.7°
SWELLING	0.95	11.5 ml	31.8 ml	42.7 ml
ATHABILITY	0.86	9.9 mm	27.4 mm	29.5 mm
ACTIVITY	1.00	—	1 task	1.9 tasks

ROMLOSS = Range of motion deficits.
SWELLING = Volumetric measurements of swelling.
ATHABILITY = Athletes' self-reports of perceived athletic ability.
ACTIVITY = Observed motor activity scores.

TABLE 3. *Intraclass correlation coefficients ($ICC_{2,1}$), standard errors of measurement (SEM), and responsiveness data.*

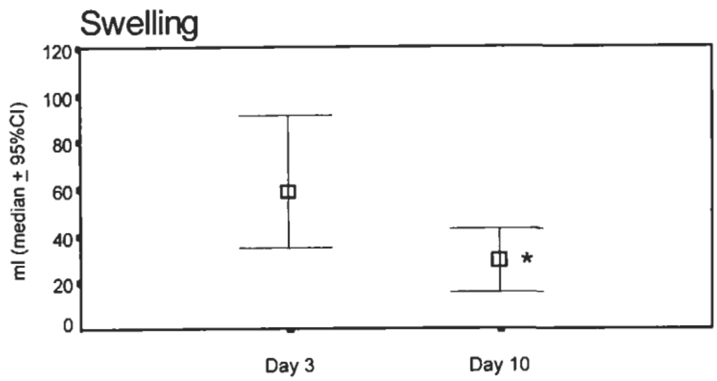

FIGURE 1. *Responsiveness of volumetric displacement (SWELLING) scores between occasions of measurement (* $p < 0.0125$; N = 13).*

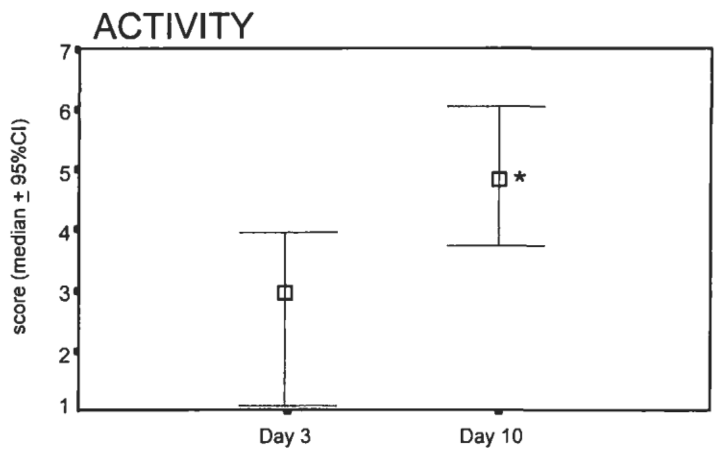

FIGURE 2. *Responsiveness of motor task completion (ACTIVITY) scores between occasions of measurement (* $p < 0.0125$; N = 13).*

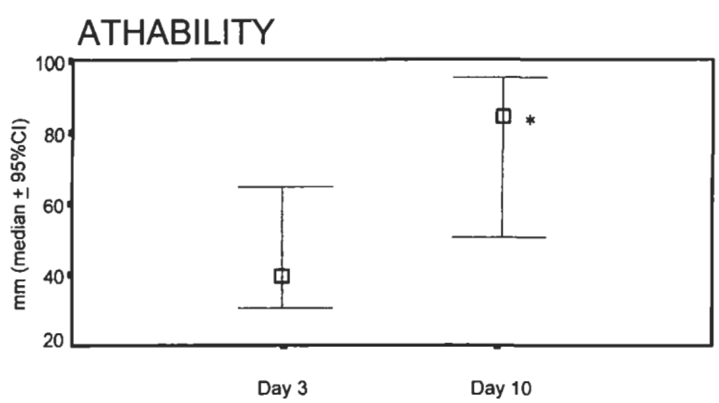

FIGURE 3. *Responsiveness of perceived athletic ability (ATHABILITY) scores between occasions of measurement (* $p < 0.0125$; N = 13).*

the minimum reliable change estimates previously calculated for each of the four indicator variables as criteria (Figure 5). Only four subjects' ROMLOSS scores (31%) demonstrated changes greater than those expected due to chance (reliable change > 1.96; $p < 0.05$) compared with six subjects for ATHABILITY (46%), seven subjects for SWELLING (54%), and eight subjects for ACTIVITY (62%).

DISCUSSION

Payton (20) has commented that one of the primary tasks of researchers is to provide clinicians with useful tools which can then be used to measure therapeutic effects. The purpose of this study was to evaluate and compare the usefulness of impairment and functional ability measures as indicators of clinical progress in a sample of collegiate athletes with acute ankle sprains. Our data indicate that sagittal plane range of motion deficits, volumetric measurement of swelling, athletes' self-reported athletic ability, and demonstrated loco- motor activity produced consistent scores within the same occasion of measurement. With the notable exception of range of motion differences measured between the involved and uninvolved extremities (ROM- LOSS), all of these measures were sensitive to changes occurring between days 3 and 10 following acute ankle sprain.

The intrarater (test-retest) reliability coefficient computed from volumetric displacement data also revealed a low proportion of measurement error to total score, confirming the high degree of precision reported using coefficients of variability in earlier studies (8,25). Intraclass correlation coefficients calculated for sagittal plane ankle range of motion deficits (ROMLOSS) in our athletes fall within the range of values previously observed among general orthopaedic patients by Elveru et al (7). When comparing reliability statistics reported in these studies, readers should note that our methods differed from those employed by Elveru et al's group in three important respects. First, observers in the earlier study did not measure ankle range of motion deficits, or ROMLOSS, as we did. Elveru et al reported the interrater reliability for measurements of ankle dorsiflexion and plantar flexion, not the difference in total range of motion between two ankles. Second, observers in the Elveru et at study were blinded to the angular measures indicated on their goniometers. This was done to eliminate observers' knowledge of their previous measurements, a potential source of intrarater bias. Finally, each of the clinicians in the Elveru et at study was free to select the anatomical landmarks used to align the goniometer. Despite these methodologic differences, the reliability statistics reported for our ROMLOSS measurements appear to support Elveru et al's claim that measurements of ankle passive range of motion appear to be fairly consistent when taken by the same therapist over a short period of time, regardless of the anatomical landmarks used to guide placement of the goniometer.

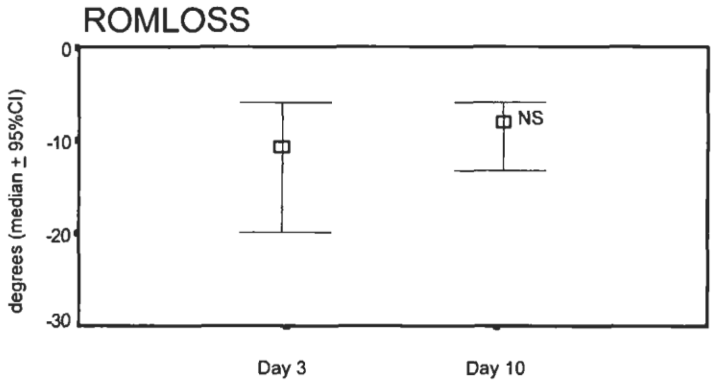

FIGURE 4. *Responsiveness of range of motion deficit (ROMLOSS) scores between occasions of measurement (NS = Nonsignificant: p > 0.0125; N = 13).*

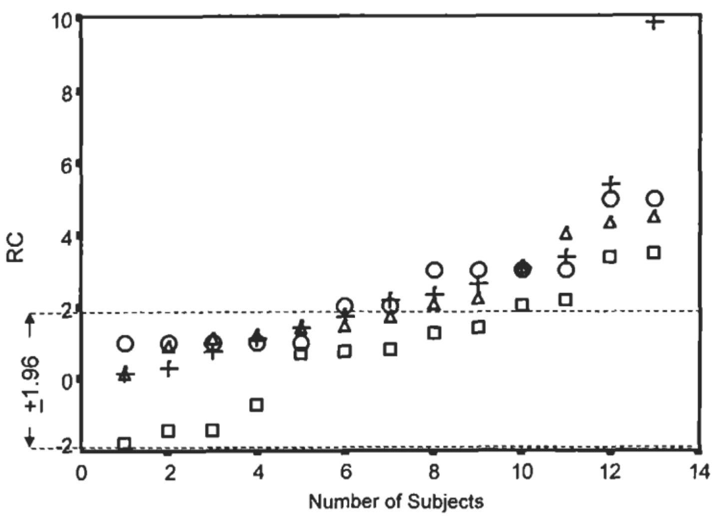

FIGURE 5. *Reliable change (RC) values of individual subjects rank ordered by measure. □ = ROMLOSS; + = SWELLING; Δ = ATHABILITY; and ○ = ACTIVITY.*

The fact that we did not gather interrater reliability data for our measures represents a weakness of the current study,. However, the nature of self-reports does not allow this type of analysis of ATHABILITY scores, and interrater agreement data for ankle range of motion deficits (7) and volumetric measures (26) are available elsewhere in the literature. We can only speculate about the degree of interobserver agreement expected for ACTIVITY scores based on this measure's demonstrated internal reliability. Kirshner and Guyatt (16) note that both the internal consistency and the interrater reliability of scores obtained from a behavioral measure are related to the amount of interpretive judgment or conjecture required on the part of the observers. Both internal consistency and interrater agreement coefficients benefit from dichotomous scoring of well-defined observable phenomena, such as motor task completion. Motor task completion was the criterion used to determine ACTIVITY scores The internal consistency of motor activity scores observed in this study (Cronbach's alpha = 0.90) indicates that relatively little observational error is contained in this measure. Consequently, we would also expect to see a high degree of interobserver agreement among subjects' motor activity scores. This hypothesis should be tested in future studies.

Range of motion measurements are commonly used to document patient progress in clinical practice and research. However, goniometric measurements of motion deficits in the sagittal plane of the ankle failed to demonstrate acceptable levels of responsiveness within our sample. Although the responsiveness of clinical measures has not been extensively studied, weak relationships between physical impairment and disability have been observed among subjects with ankle sprains in previous investigations (4,5). This finding may be related to the fact that the grade I and II ankle sprains sustained by our athletes, although severe enough to curtail athletic participation, resulted in only mild range of motion deficits. Consequently, we observed little between-subject variability in baseline ROMLOSS scores compared with the other measures employed in this study. This may account for the fact that SWELLING, ACTIVITY, and ATHABILITY scores appear to be more useful as indicators of change than ROMLOSS.

In this study, we employed a reliable change index to determine whether individual changes occurring between two occasions of measurement exceeded those expected due to chance. Although group values for SWELLING, ATHABILITY, and ACTIVITY all demonstrated statistically significant changes between occasions of measurement, a bare majority of individual scores demonstrated changes greater than would be expected due to measurement error or random variation. Our findings suggest that these measures may not reliably detect changes occurring between roughly the third and tenth day postinjury for about 40-70% of athletes with grade I and II ankle sprains. This would be particularly true for those individuals displaying relatively mild impairments or functional limitations (ie., measured value < reliable change value) during the initial clinical evaluation. We recommend that investigators consider using reliable change index values to define minimally meaningful effect sizes in future studies of ankle sprains.

Many clinicians tend to minimize the importance of self-reports and other "subjective" clinical data in clinical decision-making in the belief that such measures are "soft" and untrustworthy indicators of a patient's disability (3,22,23). It may surprise some readers to note that sim ple motor performance and self- report measures employed 3 days postinjury were as responsive to change as highly reliable impairment measures such as volumetric displacement. However, based on their demonstrated reliability and responsiveness within the context of this study, the assumption that behavioral and self-report measures are categorically less useful than physical measures should be rejected by sports medicine practitioners.

Readers should be aware that this study has certain limitations of design and implementation, which may influence the interpretation of these results. We chose to operationalize the concept of impairment by using those physical measures most frequently employed as dependent variables in previous ankle sprain studies. Range of motion and swelling were selected based on the results of our literature review and the prevalence of these measures in clinical practice settings. Although we believe these to be acceptable criteria for determining content validity, we readily acknowledge that the use of other measures to operationally define physical impairment could have led to different results in this study. We also are aware that the homemade measures of functional ability used here are unproven and unrefined. Nonetheless, we believe that they reflect a common decision-making practice. Athletes are frequently asked to perform selected locomotor activities, then rate their readiness to return to participation based on their demonstrated motor performance and symptomatic responses to these stimuli. These observable behavioral responses may be proxy variables which reflect latent personality traits, such as motivation levels and hardiness factors, which may affect an individual's decision to return to participation following an ankle sprain. Readers are cautioned that the results obtained from our small sample of collegiate athletes with acute ankle sprains may not generalize to other subjects or clinical settings. In addition, the responsiveness of measures calculated in the current study are based on observations made at an interval of only 1 week between measurements. These candidate indicators may not all change at the same rates during other periods during rehabilitation, and their relative responsiveness may also vary over longer or shorter courses of treatment.

We must continue to study new and existing clinical measures in order to determine clinically meaningful effect sizes and ensure the trustworthiness of measures used to relate therapeutic practices to functional outcomes (12,18). In order to attain these goals, future studies should examine the relationships between measures of impairment, functional limitation, and outcome in order to:

1) determine whether data gathered from athletes fit conceptual models of disablement designed for general populations (construct validity), and
2) determine the extent to which organic and behavioral measures result in accurate predictions in the number of days lost due to injury following ankle sprain (predictive validity).

CONCLUSION

This study was conducted to examine the usefulness of selected impairment and functional ability measures for the purpose of assessing treatment response following acute ankle sprains in athletes. A measure's usefulness for this purpose depends on proof of its reliability and responsiveness. That is, it must demonstrate stability during periods in which the underlying trait would be expected not to change (reliability), while also being able to detect true changes occurring over time or in response to therapy (responsiveness) (16). Our results suggest that locomotor task performance and athletes' self- reported responses to these motor activities may be at least as useful as physical measures for this purpose.

ACKNOWLEDGMENTS

The authors wish to thank our colleague, David Martin, PhD, ATC, for allowing access to subjects and use of equipment associated with a concurrent study sponsored by Electropharmacology, Inc.

REFERENCES

1. Airaksinen 0, Kolari PJ, Miettinen H: Elastic bandages and intermittent pneumatic compression for treatment of acute ankle sprains. Arch Phys Med Rehabil 71:380-383, 1990
2. Christensen L, Mendoza JL: A method of assessing change in a single subject: An alteration of the RCI index. Behav Ther 17:305-308, 1986
3. Delitto A: Subjective measures in clinical decision making. Phys Ther 69:585-589, 1989
4. Dettori JR, Basmania Q: Early ankle mobilization. Part II: A one-year follow-up of acute, lateral ankle sprains. Mil Med 159:20-24, 1994

5. Dettori JR, Pearson BD, Basmania Lednar WM: Early ankle mobilization. Part I: The immediate effect on acute, lateral ankle sprains. Mil Med 159:1520, 1994
6. Eiff MP, Smith AT, Smith GE: Early immobilization in the treatment of lateral ankle sprains. Am J Sports Med 22:8388, 1994
7. Elveru RA, Rothstein JM, Lamb RL: Goniometric reliability in a clinical setting. Subtalar and ankle joint measurements. Phys Ther 68:672-677, 1988
8. Goldie 1, Gunterberg B, Jacobson C: Foot volumetry as an objective test of the effect of antiphlogistic drugs in ankle sprains. Rheum Rehabil 13:204-207, 1988
9. Guyatt G, Walter S, Norman G: Measuring change over time: Assessing the usefulness of evaluative instruments. Chron Dis 40:171-178, 1987
10. Hocutt JE Jr, Jaffe R, Rylander CR, Beebe JK: Cryotherapy in ankle sprains. Am I Sports Med 10:316-319, 1982
11. Jacobson NS, Follette WC, Revensdorf D: Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. Behav Ther 15:336-352, 1984
12. Jette AM: Outcomes research: Shifting the dominant research paradigm in physical therapy. Phys Ther 75:965970, 1995
13. Johnston MV, Findley TW, DeLuca Katz RT: Research in physical medicine and rehabilitation. XII: Measurement tools with application to brain injury. Arch Phys Med Rehabil 70:40-56, 1991
14. Kaikkonen A, Kannus P, Jarvinen M: A performance test protocol and scoring scale for the evaluation of ankle injuries. Am J Sports Med 22:462-469, 1994
15. Kazdin AE: Single-Case Research Designs: Methods for Clinical and Applied Settings, New York, NY: Oxford University Press, Inc., 1982
16. Kirshner B, Guyatt G: A methodological framework for assessing health indices. J Chron Dis 38:27-36, 1985
17. Linde F, Hvass I, Madsen F: Compression bandage in the treatment of ankle sprains. Scand J Rehabil Med 16:177179, 1984
18. Michels E: Letter to the editor. Phys Ther 74:1138-1139, 1994
19. Pasila M, Visuri T, Sundholm A: Pulsating shortwave diathermy: Value in treatment of recent foot and ankle sprains. Arch Phys Med Rehabil 59:383-386, 1978
20. Payton OD: Measurement validity in physical therapy research (commentary). Phys Ther 73:114-115, 1993
21. Pennington GM, Danley DL, Sumko MH, Bucknell A, Nelson jH: Pulsed, non-thermal, high-frequency electromagnetic energy (DIAPULSE) in the treatment of grade I and grade II ankle sprains. Mil Med 158:101-104, 1993
22. Rothstein: On defining subjective and objective measurements. Phys Ther 69:577-579, 1989
23. Rothstein JM, Echternach JL: Primer on Measurement: An Introductory Guide to Measurement Issues, Alexandria, VA: American Physical Therapy Association, 1993
24. Shrout PE, Fleiss JL: lntraclass correlations: Uses in assessing rater reliability. Psychol Bull 86:420-428, 1979
25. Smyth CI, Velayos EE, Hlad CJ Jr: A method for measuring swelling of the hands and feet. Part I: Normal variations and applications in inflammatory joint diseases. Acta Rheum Scand 9:293-305, 1963
26. Waylett-Rendal1J, Seibly DS: A study of the accuracy of a commercially available volumeter. J Hand Ther 1:10-13, 1991

27. Wilson DH: Treatment of soft-tissue injuries by pulsed electrical energy. BMJ 2:269-270, 1972